PATENT
Attorney Docket No. GEN-001

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

APPLICANT:	Chandra et al.	CONF NO.:	5207
SERIAL NO.:	10/644,582	GROUP NO.:	2166
FILING DATE:	August 20, 2003	EXAMINER:	Pham, Khanh
TITLE:	System, Method and Apparatus for Assembling and Mining Life Science Data		

Mail Stop Appeal Briefs - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

<u>AMENDED APPEAL BRIEF</u>

This Amended Appeal Brief is submitted in accordance with 37 C.F.R. § 41.37 in furtherance of the Notice of Non-Compliant Appeal Brief issued on June 26, 2008 and the Notice of Appeal filed April 18, 2008, in support of the appeal from final rejection of pending claims in the above-identified application.

Appellant believes that no extension fees are due for this Appeal Brief to be entered and considered. However, please consider this a conditional petition for the proper extension, if one is required. The Commissioner is hereby authorized to charge any additional fees that may be due, for further extensions of time or any other purpose associated with this submission, or credit any overpayment, to Appellant's undersigned counsel's deposit account number 07-1700 with reference to docket number GEN-001.

REAL PARTY IN INTEREST

The real party in interest is the owner of the present application, Genstruct, Inc.

RELATED APPEALS AND INTERFERENCES

No other appeals or interferences directly affect or will be directly affected by the Board's decision in the present appeal.

STATUS OF CLAIMS

The application as filed contained 91 claims. Applicants subsequently cancelled claims 41-91. A restriction requirement issued and in response Applicants elected claims 26-40 and withdrew claims 1-25 from consideration. In an amendment filed on July 5, 2006, Applicants amended claims 26 and 34-37 and and added new claims 92-106. In an amendment filed on February 28, 2007 Applicants amended claims 26, 37, 98, 99 and 101. On November 9, 2007, Applicants amended claim 26. Claims 26-40 and 92-106 remain pending, have been rejected, and are the subject of this appeal.

STATUS OF AMENDMENTS

An amendment is being filed concurrently herewith in order to simplify the issues on appeal. The Claims Appendix presumes entry of this amendment.

SUMMARY OF CLAIMED SUBJECT MATTER

The present invention involves the storage of life science data.[1] Unlike conventional databases, the invention stores the life science data in a format that exposes the relationships within the information, facilitates efficient knowledge mining, and makes the information more readily comprehensible and available.[2]

In one aspect and according to claim 26, the invention relates to a system for storing life science information that includes an electronic database storage module and an inference engine.[3] The electronic database storage module stores a library of case frames.[4] Each case frame includes at least two unspecified object identifiers and a relationship connector, wherein the relationship connector relates the identifiers to each other based on a causal relationship between the object identifiers and is based on a life science ontology.[5] The inference engine module operates on the case frames to manage the addition of new data to a database of life science data.[6] In so doing, the inference engine translates data to be added to the database into a form compatible with the database.[7] One of the case frames is selected as a template to represent the new data based on the life science ontology.[8] Elements of the new data are assigned to the object identifiers and relationship connectors, thus assuring newly created life science assertions conform to the life science ontology.[9]

---

[1] Specification at paragraph [0009].
[2] Specification at paragraph [0010].
[3] Specification at paragraph [0167].
[4] Specification at paragraph [0068].
[5] Specification at paragraph [0020].
[6] Specification at paragraph [0172].
[7] Specification at paragraph [0155].
[8] Specification at paragraph [0151].
[9] Specification at paragraph [0017],

3

In another aspect and according to claim 101, the invention relates to a system for storing life science data including an electronic database storage device for storing case frames.[10] Each case frame includes at least two object identifiers and a relationship connector.[11] The relationship connector relates two of the at least two object identifiers to each other and is based on a life science ontology.[12] The database includes case frames representing at least enzyme reactions, binding interactions, modifications of polymers, protein phosphorylation reactions, gene expressions, acetylation, peptide-bond cleavage, glycosylayion, lipidation, fatty-acylation, methylation, metallation, cross-linking, hydroxylation, sulfation, ADP-ribosylation, translocation, and transcriptional activations.[13]

GROUNDS FOR REJECTION TO BE REVIEWED ON APPEAL

The issues on appeal are whether claims 26-39, 92-95 and 97-106 are unpatentable under 35 U.S.C. §102(e) over U.S. Patent Publication Number 2002/0087275 to Kim et al. ("Kim"), and whetiner claims 40, 96 and 99-100 are unpatentable under 35 U.S.C. §103(a) in light of Kim and further in view of U.S. Patent Publication No. 2002/0198858 to Stanely, et al. ("Stanley").

---

[10] Specification at paragraphs [0010] and [0068].
[11] Specification at paragraph [0020].
[12] Specification at paragraph [0020].
[13] Specificaiton at paragraph [0130].

ARGUMENT

The Invention

Traditional methods of discerning and understanding the meaning of life science information do not sufficiently address the large amount of material that must be absorbed and combined. New and old information are presented and stored in a variety of sources and formats, different languages, different data structures, conflicting vocabulary and ontology, and often are presented based on inconsistent and competing theories. To form an effective understanding of a biological system, a life science researcher must synthesize information from many of these sources.

Understanding a biological system may require that information of many different types and at many different levels be combined. Life science information may include material on basic chemistry, proteins, cells, tissues, and effects on organisms or population – all of which may be interrelated. These interrelations may be complex, poorly understood, or hidden. Furthermore, the exponential growth in the amount of biological data presents a cognitive barier to understanding biological mechanisms, as the sheer amount of data relevant to a particular mechanistic hypothesis is beyond the reasoning capacity of scientists. An analysis conducted using Appellants' knowledgebase and the claimed technology may involve consideration of data from thousands of scientific papers, and new biological mechanistic hypotheses developed from such analyses often include data from hundreds of papers or more.

Applicant's invention thus provides a system that makes biological information available in a manner that enhances understanding of the interrelationships within the

information by facilitating the assembly and mining of life science data in a comprehensive and structured manner. As a result, the biological data is structured in the system to result in production of a knowledgebase that is an improvement over conventional databases as it is augmented with content rationalizing relationships among the data and thus permitting automated reasoning. A library of case frame templates are used to translate biological data into a form compatible with an existing database of life science data. Each case frame template includes a predefined collection of relationship connectors between unspecified objects with defined object types. The connectors are based on a causal relationship between the objects and are based on a consistent life science ontology.

When new information is added to the database, an inference engine selects an appropriate case frame template to use to represent the new data and creates a case frame from the template by assigning elements of the new data to previously unspecified object identifiers in the case frame. As a result, new life science assertions are created in the database that are consistent with the ontology. The cited reference does not utilize case frame templates to create new case frames, nor does it describe the use of case frames to ensure consistency with a life science ontology.

In a more specific instance of the invention, the database includes case frames directed at specific life science concepts, including enzyme reactions, binding interactions, modifications of polymers, protein phosphorylation reactions, gene expressions, acetylation, peptide-bond cleavage, glycosylayion, lipidation, fatty-acylation, methylation, metallation, cross-linking, hydroxylation, sulfation, ADP-

ribosylation, translocation, and transcriptional activations. Neither cited reference describes a database containing such a comprehensive collection of entities.

The Examiner's Rejection of Claim 26 Under 35 U.S.C. §102(e)

In rejecting the independent claim claim 26, the Examiner essentially ignores a key element of the claim. Independent claim 26 recites an inference engine that manages the addition of new data into the database by, in part, selecting a case frame template to represent the new data. The cited reference neither teaches nor suggests this feature.

The Examiner suggests "Kim also teaches at Tables 1-3 on pages 6-7 three different templates."[14] These "tables" fail to anticipate the claims on two accounts. First, the structure of the disclosed tables do not contain all the components of the claimed case frames. Second, nowhere in Kim is there any indication that a particular case frame is "selected."

---

[14] Final Office Action at pg. 10.

A Table is not a Case Frame

The tables cited by the Examiner are just that – a "set of (property_name; value) pairs" that together provide descriptive information about "a vertex."[15] None of the three tables illustrated by Kim include information about more than one biological entity, let alone the relationships among the entities. A simple listing of property values (e.g., chromosome name, location, function etc.) does not provide any information about how that particular biological entity interacts with other entities, how it fits within a larger system or how its regulation influences (or is influenced by) other entities. While providing tabular data entry mechisms may provide a basic data input structure for users to enter data into the Kim system, it does not facilitate the programmatic ingestion of large amounts of biological data *in toto* as do the case frames of the claimed invention.

Without the aide of templates that include "at least two" object identifiers and "a relationship connector" the entry of information into the Kim system must be done on an entity-by-entity basis, with a separate step of "connecting" vertices after the fact. More specifically, Kim relies on a two-step method for creating data in the database. First, Kim details a series of steps for creating individual vertices that represent yeast genes.[16] Kim then describes a subsequent step of defining edges that require the definition of various components (including the vertices) as prerequisites.[17] In practice, no matter how much "data" is entered into the Kim system (whether entry be automated or manual) the result is simply a collection of tables, each having to be built and populated

---

[15] Kim, paragraph [0077].
[16] Kim, paragraphs [0080] – [0100].
[17] Kim, paragraph [0095].

independently of each other. Such a result must then rely on human intervention to determine how the entities relate to each other and then define the "edges."

The claimed invention reduces, or in some cases even eliminates the need for manual processes by using the pre-constructed case frame templates that include the ability to have the system automatically instantiate a model of a biological system having "two or more" components and predefined relationships among the components.

One of the key functions of the case frame templates is to define rules and semantics describing what life science data can be connected and how to connect it. This structure addresses the challenge of creating a large database of biological concepts generated by multiple scientists over many years and ensuring the consistency of how information is added to the database. For example, a tissue type cannot be a reactant in a reaction. This may seem like a trivial statement, but maintaining a strict case frame structure implying hundreds of such restrictions has enabled the efficient creation of a coherent life science database from over 40,000 different sources.

No Selection of Case Frames

Even if one were to equate Kim's tables with the more comprehensive case frames (which, as illustrated above is not appropriate), Kim still fails to anticipate the claims because there is no mention that the Kim system "selects one of the case frames as a template to represent the new data."[18] By maintaining a library of pre-defined case frames that include both object type information and pre-defined connectors that are based on causal relationships between the object identifiers, an inference engine can

select an appropriate case frame through which to add new data into the database that includes both the object identifiers and the relationships among the identifiers such that new life science assertions added to the database conform to a specified life science ontology. In doing so, Applicants' invention facilitates an automated and rapid expansion of the database in a manner that is beyond the manual means of first creating vertices and subsequently defining edges between vertices as described by Kim.

As such, the case frames go well beyond simple data structures described in Kim and actually enforce the structure and biological relationships of a defined life science ontology. Using this data assertion process, the Assignee has, as of this date, built a database comprising millions of assertions, using only a handful of curators to do so. In contrast, the Kim approach would require countless hours of manual labor to build tables for each biological entity, to populate the tables, and to identify relationships among the entities.

The burden is on the Examiner to demonstrate that each feature of a claim is met by a reference or valid combination of references. The courts have repeatedly and consistently held that "all limitations [of a claim] must be considered ... and it is error to ignore specific limitations in distinguishing over the references." *In re Boe and Duke*, 184 USPQ 38, 40 (CCPA 1974); *see also* MPEP §2143.03. In the present case, the reference cited by the Examiner does not realize the claimed invention. The deficiencies of Kim cannot be cured by Stanley because, like Kim, it neither teaches nor suggests the structure or selection of case frames as claimed.

---

[18] Independent Claim 26

These differences are not only core to the invention, but the successful implementation of the invention. Systems like Kim and Stanley have been used to model small biological systems and protein groups for some time, but to applicants' knowledge have never been able to scale to the extent necessary to study large complex biological systems. It is precisely the creation of the claimed case frames and the ability to select a case frame as a template for new biological data that has made the invention commercially viable. Without the ability to rapidly ingest large amounts of biologically diverse data into the database, comprehensive studies become incomplete and of little value to large pharmaceuticals.

In contrast, the knowledgebase created using the claimed techniques has become an invaluable aide in the drug development process. Specifically, the inventors of this technology have to date created a knowledgebase of millions of nodes derived from life science databases and over 40,000 reference sources. Using this novel knowledgebase, the inventors have performed over 40 successful commercial projects, enabling compounds to begin clinical trials, developing biomarkers used to assess the efficacy of marketed compounds, and defining a mechanism that in some cases have been a centerpiece of a new drug application.

The Examiner's Rejection of Claim 101 Under 35 U.S.C. §102(e)

In rejecting claim 101, the Examiner relies on an inherency argument to reject the claim, while failing to specifically cite to any references or provide any extrinsic evidence. Independent claim 101 recites, in part, a system for storing life science data that includes an electronic database storage device for storing case frames wherein the

11

database comprises case frames representing at least enzyme reactions, binding

interactions, modifications of polymers, protein phosphorylation reactions, gene

expressions, acetylation, peptide-bond cleavage, glycosylayion, lipidation, fatty-

acylation, methylation, metallation, cross-linking, hydroxylation, sulfation, ADP-

ribosylation, translocation and transcriptional activations.

To establish inherency, the extrinsic evidence "must make clear that the missing

descriptive matter is necessarily present in the thing described in the reference, and that it

would be so recognized by persons of ordinary skill." Continental Can Co. v. Monsanto

Co., 948 F.2d 1264, 1268, 20 USPQ2d 1746, 1749 (Fed. Cir. 1991). Moreover, the basis

for inherency must be clear; the Federal Circuit has repeatedly held that the relevance of

a reference cannot be predicated on "mere conjecture." In re Robinson W.L. Gore &

Associates, Inc. v. Garlock, Inc., 721 F.2d 1540, 220 U.S.P.Q. 303 (Fed. Cir. 1983), cert.

denied, 469 U.S. 851, 105 S.Ct. 172 (1984); Continental Can, supra, 948 F.2d at 1268, 20

USPQ2d at 1749 (Fed. Cir. 1991). See also In re Robertson, 49 USPQ2d 1949 (1999).

In the present case, not only has the Examiner failed to satisfy the stringent criteria for an

inherency rejection; he has not even cited anything relevant to the claims he has rejected.

The Examiner's Rejection of Claims 40, 96, 99 and 100 Under 35 U.S.C. §103(a)

Claims 40, 96, 99 and 100 depend either directly or indirectly from claim 26, and

are therefore patentable for the reasons discussed above.

CONCLUSION

For all of the foregoing reasons, we submit that the Examiner's rejections of

claims 26-40 and 92-106 were erroneous, and reversal thereof is respectfully requested.

Accompanying this brief is the fee specified in 37 C.F.R. §1.17(f). Please charge

any additional fee occasioned by this paper to our Deposit Account No. 07-1700.

Respectfully submitted,

Date: June 27, 2008

Electronic Signature: /Joel E. Lehrer/
Joel E. Lehrer
Reg. No. 56,401
Goodwin Procter LLP
Exchange Place
Boston, MA 02109
Tel. No. (617) 570-1057
Fax No. (617) 523-1231

CLAIMS APPENDIX

26. A system for storing life science information, the system comprising:

an electronic database storage module for storing a library of case frames, each case frame comprising:

at least two unspecified object identifiers;

a relationship connector, wherein the relationship connector relates two of the at least two object identifiers to each other based on a causal relationship between the object identifiers and is based on a life science ontology; and

an inference engine module for:

managing the addition of new data to a database of life science data by translating the new data into a form compatible with the database;

selecting one of the case frames as a template to represent the new data based at least in part on the life science ontology; and

assigning elements of the new data to the object identifiers and relationship connectors, thus assuring newly created life science assertions conform to the life science ontology.

27. The system of claim 26, wherein a set of said case frames define a biological function.

28. The system of claim 27, wherein the biological function comprises a chemical reaction.

29. The system of claim 27, wherein the biological function comprises transport.

30. The system of claim 27, wherein the biological function comprises digestion of a biomolecule.

31. The system of claim 26, wherein at least one of the at least two object identifiers identifies a biomolecule.

32. The system of claim 26, wherein at least one of the at least two object identifiers identifies a biological function.

33. The system of claim 26, wherein at least one of the at least two object identifiers identifies a relationship connector.

34. The system of claim 26, wherein the relationship connector represents an identity relationship.

35. The system of claim 26, wherein the relationship connector represents a product relationship.

36. The system of claim 26, wherein the relationship connector represents a substrate relationship.

37. The system of claim 26, wherein the relationship connector represents an enzymatic relationship.

38. The system of claim 26 further comprising a graphical user interface configured to permit a user to query the database based on the relationship connector.

39. The system of claim 26 further comprising a data input interface configured to accept user instructions relating to the creation of a new case frame.

40. The system of claim 26 further comprising an access manager configured to restrict access to one or more portions of the electronic database.

41. – 91. (Cancelled)

92. The system of claim 26 wherein the inference engine further modifies the selected case frames such that the selected case frames more accurately represent the new data.

93. The system of claim 92 wherein the modifications comprise one or more of the addition of new fields, the addition of new relationships, and the addition of metadata.

94. The system of claim 93 wherein the metadata comprises one or more of the source of the new data, the date the new data was received, the time the new data was received, and the experimental conditions under which the new data was created.

95. The system of claim 26 further comprising a harmonization and transfer module for interfacing with multiple disparate sources of life science data and receiving the new data.

16

96. The system of claim 95 wherein the received data is received in XML format.

97. The system of claim 95 wherein the harmonization and transfer module further translates the received data into a data format compatible with the case frames.

98. The system of claim 26 further comprising a discovery environment for displaying pathways among the plurality of case frames, the pathways representing causal relationships among the case frames.

99. The system of claim 26 further comprising a managed account interface for attributing access restrictions to one or more case frames in the database.

100. The system of claim 99 wherein the access restrictions comprise one or more of public access rights, subscription-based access rights, and proprietary access rights.

101. A system for storing life science data, the system comprising:

an electronic database storage device for storing a plurality of case frames, each case frame comprising:

at least two object identifiers; and

a relationship connector, wherein the relationship connector relates two of the at least two object identifiers to each other and is based on a life science ontology; and

wherein the database comprises case frames representing at least enzyme reactions, binding interactions, modifications of polymers, protein phosphorylation reactions, gene expressions, acetylation, peptide-bond cleavage, glycosylayion, lipidation, fatty-acylation, methylation, metallation, cross-linking, hydroxylation, sulfation, ADP-ribosylation, translocation and transcriptional activations.

17

102. The system of claim 101 wherein the case frame representing protein phosphorylation reactions comprises a reactant, a product, and a catalyst.

103. The system of claim 101 wherein the case frame representing gene expressions comprises a gene and a gene product.

104. The system of claim 101 wherein the case frame representing transcriptional activation comprises a gene expression, an activation, and a transcriptional activator.

105. The system of claim 101 further comprising a harmonization and transfer module for interfacing with multiple disparate sources of life science data and receiving new data for inclusion in the database.

106. The system of claim 105 further comprising an inference engine for managing the addition of the new data by instantiating a subset of the plurality of case frames to represent the new data and assuring the instantiated case frames conform to the life science ontology, thereby creating life science assertions in the database.

EVIDENCE APPENDIX

There has been no evidence submitted under 37 C.F.R. 1.130, 1.131 or 1.32 in this case.

RELATED PROCEEDINGS APPENDIX

There have been no proceedings related to this case.

LIBC/3329334.1